Smart Energy Policies for Sustainable Mobile Networks via Forecasting and Adaptive Control

Ángel Fernández Gambín and **Michele Rossi**
Department of Information Engineering, University of Padova
via G. Gradenigo, 6/B, 35131 Padova, Italy
{afgambin, rossi}@dei.unipd.it

***Abstract*—The design of sustainable mobile networks is key to reduce their impact on the environment, and to diminish their operating cost. As a solution to this, we advocate Energy Harvesting (EH) Base Stations (BSs) that collect energy from the environment, use it to serve the local traffic and/or store it in a battery for later use. Moreover, whenever the amount of energy harvested is insufficient to serve their traffic load, BSs purchase energy from the power grid. Within this setup, a smart energy management strategy is devised with the goal of diminishing the cost incurred in the energy purchases. This is achieved by intelligently controlling the amount of energy that BSs buy from the electrical grid over time, by accounting for the harvested energy, the traffic load, and hourly energy prices. The proposed optimization framework combines pattern forecasting and adaptive control. In a first stage, harvested energy and traffic load processes are modeled through a Long Short-Term Memory (LSTM) neural network, allowing each BS to independently predict future energy and load patterns. LSTM-based forecasts are then fed into an adaptive control block, where foresighted optimization is performed using Model Predictive Control (MPC). Numerical results, obtained with real-world energy and load signals, show cost savings close to 20% and reductions in the amount of energy purchased from the electrical grid of about 24%, with respect to a heuristic scheme where future system states are not taken into account.**

***Index Terms*—Energy management, energy harvesting, forecasting, adaptive control, energy self-sustainability, mobile networks.**

I. Introduction

The large use of Information and Communications Technologies (ICT) is increasing the amount of energy drained by the telecommunication infrastructure and hence its footprint on the environment. As an example, ICT account for 8-10$\%$ of the European electricity consumption and up to 4% of its carbon emissions [1]. Moreover, forecasts for 2030 are that 51% of the global electricity consumption and 23% of the carbon footprint by human activity will be due to ICT [2]. These figures are going to worsen due to the increasing trend in the traffic demand, and in the number of connected devices, especially mobile [3]. Besides its environmental impact, the energy drained to operate ICT infrastructures largely affects the revenue of mobile network operators. Therefore, the design of sustainable mobile networks is a priority for any future development in the ICT sector.

In this work, we advocate sustainable mobile networks deployments where BSs are equipped with EH hardware, namely, solar panels and energy storage devices. Base stations collect energy from the environment, store it and use it to serve their local traffic demand over time. Whenever the harvested energy is insufficient to seve the traffic load, BSs can purchase energy from the power grid.

A large body of work has recently appeared on the use of energy harvesting in mobile networks. A base station power management scheme is presented in [4] considering exogenous processes such as: renewable power generation, power price, and wireless traffic load. The scenario is modeled as a stochastic optimization problem and then transformed into a linear programming one. However, the exogenous processes are modeled through simplified models, i.e., as stochastic time series with values defined within finite state sets, which fail to faithfully capture the actual system behavior. The authors of [5] consider BSs with energy harvesting capabilities connected to the power grid as a means to carry out the energy trading. A joint optimization tackling BS operation and power distribution is performed to minimize the on-grid power consumption of the BSs. Moreover, wired energy transfer to/from the power distribution network, and a user-BS association scheme based on cell zooming are investigated. The problem is split into two subproblems, which are solved through heuristics. A similar approach is considered in [6], where two problems are considered: the first one consists of optimizing the energy allocation at individual BSs to accommodate for the temporal dynamics of harvested energy and mobile traffic. Considering the spatial diversity of mobile traffic patterns, the second problem's objective is to balance the energy consumption among BSs, by adapting the cell size (radio coverage) in an attempt to reduce the on-grid energy consumption of the mobile network. Again, the solutions are obtained through heuristic algorithms. An optimal energy management strategy to minimize the energy cost incurred by a set of cellular base stations is presented in [7]. In this paper, base stations can exchange energy with the power grid and are equipped with batteries (energy storage) and renewable energy harvesting devices. Simulation results show that a cost reduction can be achieved by increasing the battery capacity of each BS and/or the number of base stations.

Paper contribution: in the present work, we present a smart energy management framework with the objective of diminishing the monetary cost incurred in purchasing energy from the power grid. This is achieved by controlling the amount of energy that BSs buy from the electrical grid over time, considering the energy that they harvest, the traffic load, and hourly energy prices (set by the power grid oper-

978-1-5386-4920-6/18/$31.00 ©2018 IEEE

ator). The optimization process employs pattern forecasting and adaptive control techniques. Specifically, the harvested energy and traffic load processes are predicted through a Long Short-Term Memory (LSTM) neural network, allowing each BS to independently track its own energy and load patterns. These forecasts are then fed into a subsequent adaptive control phase, where foresighted optimization is performed through MPC. The combined use of prediction and optimization makes it possible to make informed decisions, which take the future system evolution into account. Numerical results, obtained with real-world harvested energy traces, traffic load patterns, and power energy prices show that the proposed strategy achieves cost savings close to 20% and reduces the amount of energy purchased from the power grid of about 24% with respect to energy management policies that do not take into consideration future system states.

Paper novelty: despite the existence of previous papers on this topic, such as those aforementioned, in this paper we consider a more realistic setup, where the energy harvesting process and the BS traffic load are unknown and fully stochastic, and come from real-world traces. Also, although pattern forecasting has already been used in combination with control theory, to the best of our knowledge, this is the first time where the combination of LSTM neural networks and MPC-based adaptive control is used in an energy harvesting mobile network scenario. Lastly, the overall optimization framework that we put forward uses online and convex tools. This limits the complexity and computation time of the energy scheduler, making it applicable to real-world cases.

The paper is organized as follows. In Section II, we describe the network scenario. The optimization framework is described in Section III, where the two stages are presented: pattern forecasting in Section III-A and adaptive control in Section III-B. The numerical results are presented in section IV, and final remarks are given in section V.

II. System Model

We consider a mobile network setup comprising N EH BSs, where each BS is equipped with a solar panel, an energy conversion module and an energy storage device. All BSs can also obtain (purchase) energy from the electrical grid. The network is controlled in a time-slotted fashion, where the time slot duration Δ is fixed, but can be tuned to meet energy and load dynamics. In this work, the time granularity of the control is set to $\Delta = 1$ hour.

A. Harvested Energy Traces

The amount of solar energy harvested by each BS $n = 1, \ldots, N$ in time slot t, denoted by $H_n(t)$, has been obtained through the SolarStat tool [8]. For the solar modules, the commercially available Panasonic N235B photovoltaic technology is considered, and each solar panel has an area of $0.44\,\mathrm{m}^2$, which is deemed practical for installation in a urban environment. An example of energy harvested trace is shown in Fig. 1.

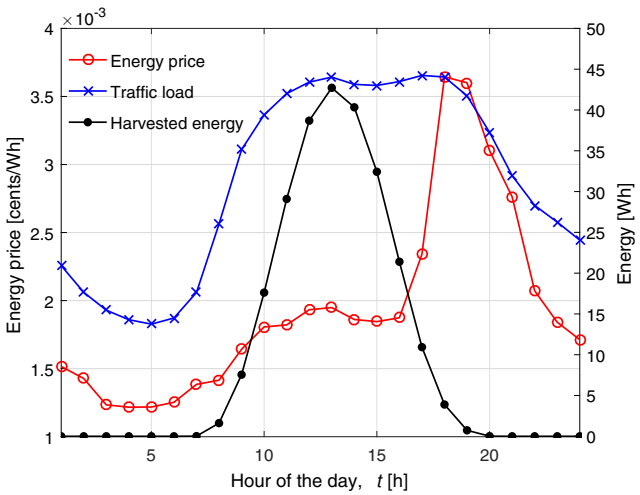


Fig. 1: Traces example: harvested energy (black curve), BS traffic load (blue curve) and energy price (red curve).

B. Base Station Energy Consumption

Traffic load traces have been obtained using real mobile data from the Big Data Challenge organized by Telecom Italia Mobile (TIM) [9]. The dataset is the result of a computation over the Call Detail Records (CDRs), logging the user activity within the TIM cellular network for the city of Milan during the months of November and December 2013. For the traffic load traces we use the CDRs related to SMS, calls and Internet activities, performing spatial and temporal aggregation. In this way, we obtain a daily traffic load profile for each BS (see, e.g., the blue curve in Fig. 1). Clustering techniques have been applied to the TIM dataset to capture the behavior of the mobile data. Specifically, the X-means algorithm [10] is used to classify the load profiles into several categories. In our numerical results, each BS $n = 1, \ldots, N$ has an associated load profile $L_n(t)$, representing the percentage of the total bandwidth that BS n allocates to serve the users in its radio cell, which is picked at random as one of the five possible clusters identified from the TIM dataset. Given the normalized load $L_n(t) \in [0, 1]$, the BS energy consumption within a time slot, $O_n(t)$, is computed using the linear model of [11]: $O_n(t) = \Delta \times (P_0 + \alpha L_n(t))$, where Δ is the time slot duration, P_0 [W] is the *load independent* BS power term and $\alpha > 0$ depends on the BS type.

C. Electric Retail Pricing

Hourly electric supply charges have been taken from the US National Grid database [12], considering the energy cost for the state of New York between January 2015 and August 2017. The price that BS n has to pay to purchase energy at a certain time slot t is denoted by $p(t)$. An example energy price trace is shown in Fig. 1, over a full day.

D. Energy Storage Devices

Energy storage devices are interchangeably referred to as Energy Buffers (EBs). The EB level for BS $n = 1, \ldots, N$ is denoted by $B_n(t)$. A *reference* energy threshold is defined as

B_{ref}, with $0 < B_{\text{ref}} < B_{\max}$ (where $B_{\max}$ is the EB capacity). B_{ref} is used in the benchmark strategy that we compare against the proposed energy management scheme (more details are provided in Section IV). For a certain BS n, if t is the current time slot, the buffer level is updated at the beginning of the next time slot $t+1$ as:

$$B_n(t+1) = B_n(t) + H_n(t) - O_n(t) + U_n(t)\,, \qquad (1)$$

where $U_n(t)$ represents the energy purchased by BS n from the power grid during time slot t. $B_n(t)$ is the EB level at the beginning of time slot t, whereas $H_n(t)$ and $O_n(t)$ respectively represent the amount of energy harvested and the energy that is locally drained (to serve the connected users). Note that $B_n(t)$ is updated at the beginning of time slot t, whereas $H_n(t)$ and $O_n(t)$ are only known at the end of it. The expected behavior $\mathbb{E}[H_n(t) - O_n(t)]$ is obtained through the theory in Section III-A to make foresighted decisions, where $\mathbb{E}[\cdot]$ is the expectation operator.

III. Optimization Framework

Next, we propose an energy management framework with the objective of decreasing the monetary cost incurred by energy purchases from the power grid. This is achieved by controlling the amount of energy that BSs buy from the electrical grid over time considering harvested energy, traffic load and energy prices. The framework consists of two main blocks: (i) pattern forecasting and (ii) adaptive control. In the first block, the harvested energy and traffic load processes are predicted through a machine learning approach, specifically a Recurrent artificial Neural Network (RNN) [13]. The details of this RNN are provided in Section III-A. This allows each BS to independently track its own energy and load processes, capturing their statistical behavior and obtaining forecasts for the corresponding time series. Note that energy prices are available one-day ahead, thus their forecasting is not needed. Energy and load forecasts are then fed into the foresighted optimization approach of Section III-B. Their use allows for making informed decisions, which take the future system evolution into account. This results in an effective energy management scheme, that reduces the amount of energy that has to be purchased from the power grid. In the second block, the BS system is controlled using MPC, by determining the amount of energy $U_n(t)$ that each BS n has to purchase from the electrical grid at each time slot t. The MPC block takes online actions, considering not only the current system state, i.e., harvested energy, traffic load, energy price and EB levels, but also future ones (based on the forecasts from the forecasting block), anticipating events and acting accordingly. More details about the MPC block are provided in Section III-B.

A. Pattern Forecasting

An LSTM neural network [13] has been used to forecast the harvested energy $H(t)$ and the traffic load $L(t)$ profiles for each BS. The neurons in the hidden layers of this LSTM are Memory Cells (MCs). A MC has the ability to store or forget information about past network states by using structures called *gates*, which consist of a cascade of a neuron with sigmoidal activation function and a pointwise multiplication block. Thanks to this architecture, the output of each memory cell possibly depends on the entire sequence of past states, making LSTMs suitable for processing time series with long time dependencies [14]. An example LSTM memory cell is presented in Fig. 2. The input gate is a neuron with sigmoidal activation function (σ). Its output determines the fraction of the MC input that is fed to the cell state block. Similarly, the forget gate processes the information that is recurrently fed back into the cell state block. The output gate, instead, determines the fraction of the cell state output that is to be used as output of the MC at each time step. Gate neurons usually have sigmoidal activation functions (σ), while the input and cell state use the hyperbolic tangent ($\tanh$) activation function. All the internal connections of the MC have unitary weight [14].

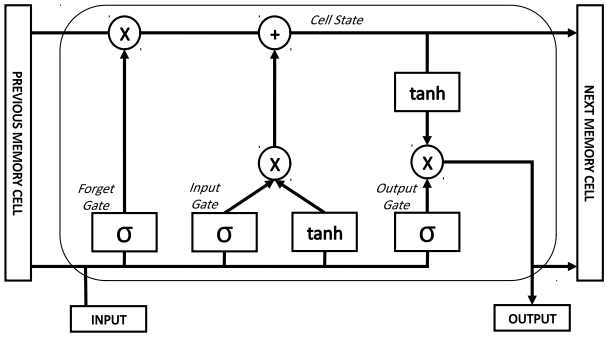
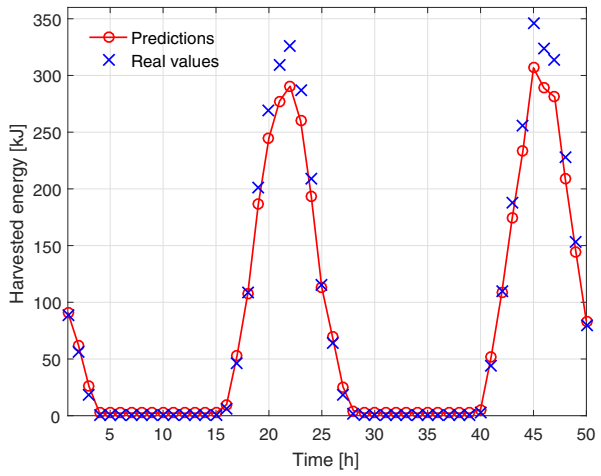


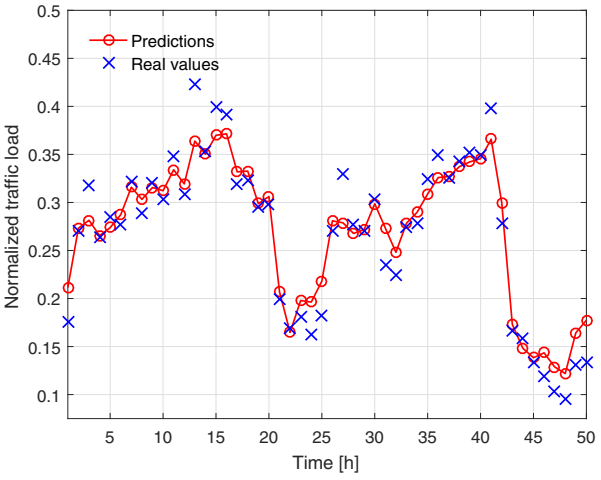
Fig. 2: LSTM memory cell diagram.

Experimental setup: to assess the performance of the proposed forecasting approach, we utilize the datasets introduced in Sections II-A and II-B. The solar generation dataset contains 175,200 samples taken every hour during a period of 20 years in the city of Los Angeles, US. The traffic load dataset has 1,440 samples taken every hour for a period of two months in the city of Milan, Italy. For both datasets, we perform M-steps ahead forecasts, where M is the optimization horizon (we use $M = 24$ hours), using sequences of 24 past time samples. 80% of the dataset is used for training, while the remaining samples are used to evaluate the accuracy of the obtained forecasts. The LSTM network has one hidden layer composed of 16 memory cells and one output neuron with linear activation function. It has been trained using the ADAM optimizer [15] and the mean squared error as the objective function. The whole setup has been implemented in Python using the high-level neural-networks-API Keras [16], running on top of Tensorflow [17].

Fig. 3 shows a prediction example. Harvested energy forecasts ($H(t)$ in Fig. 3a) are more precise than load forecasts ($L(t)$ in Fig. 3b), as solar energy traces follow quite regular bell-shape patterns, which are easier to predict. This fact is confirmed by the average Root Mean Square Error (RMSE) over the test samples: in the case of $H(t)$ we obtain RMSE $= 0.033$, whereas for $L(t)$ we get RMSE $= 0.08$.



(a) Harvested energy profile $H(t)$ forecast.



(b) Traffic load $L(t)$ forecast.

Fig. 3: LSTM forecasting examples.

B. Adaptive Control

A general MPC framework is composed of (i) an input section, (ii) an MPC controller and (iii) a real system [18]. The first block contains the prediction model (see Section III-A). The MPC solves a control problem at runtime (see below). Finally, the real system block receives the optimal actions from the MPC controller and behaves accordingly.

Notation: the system to be controlled is described through the following discrete-time model:

$$\boldsymbol{B}_{t+1} = \boldsymbol{B}_t + \boldsymbol{U}_t + \boldsymbol{W}_t, \tag{2}$$

where t is the current time slot. The $M \times N$ matrix $\boldsymbol{B}_t$ with elements b_{kn} denotes the *system state*, representing for each BS $n = 1, \ldots, N$ the energy buffer level for time slots $k = t, t+1, \ldots, t+M-1$, were $M = 24$ hours is the optimization horizon. The $M \times N$ matrix $\boldsymbol{U}_t$ with elements u_{kn} denotes the *control* matrix, representing the amount of energy that each BS n shall purchase in time slot $k = t, \ldots, t+M-1$. The $M \times N$ matrix $\boldsymbol{W}_t$ with elements w_{kn} models the *system disturbance*, i.e., the stochastic behavior of the forecast profiles (harvested and consumed energy).

Eq. (2) relates to the problem setup of Section II-D as follows: symbol $\boldsymbol{B}_t = [b_{kn}]$ contains the buffer state for all BSs, i.e., $b_{kn} = B_n(k)$, $\boldsymbol{U}_t = [u_{kn}]$ is the control, which corresponds to the amount of energy that BS n must purchase in time slot $k = t, t+1, \ldots, t+M-1$, i.e., $u_{kn} = U_n(k)$, and $\boldsymbol{W}_t = [w_{kn}]$ contains the exogenous processes, i.e., $w_{kn} = H_n(k) - O_n(k)$, which are predicted through the framework of Section III-A.

In addition to the previous variables, we define an $M \times N$ energy demand matrix, $\boldsymbol{D}_t = [d_{kn}]$, where d_{kn} represents the power grid energy that BS n needs to serve the expected traffic load in time slot k. It is calculated as follows:

$$\begin{cases} d_{kn} = 0 \quad if \quad w_{kn} \geq 0, \\ d_{kn} = |w_{kn}| \quad if \quad w_{kn} < 0. \end{cases} \tag{3}$$

Objective function: the goal of the MPC controller is to determine the amount u_{kn} that each BS n should purchase from the electrical grid in time slot $k = t, \ldots, t+M-1$. The following cost function tracks the total amount of energy purchased by the BSs ($p(k)$ is the energy price at time k):

$$f_{\text{MPC}}(\boldsymbol{U}_t, \boldsymbol{p(t)}) = \sum_{k=t}^{t+M-1} \sum_{n=1}^{N} p(k) u_{kn}. \tag{4}$$

Control problem: the following finite-horizon multi-objective optimization problem is formulated:

$$\min_{\boldsymbol{U}_t} \quad f_{\text{MPC}}(\boldsymbol{U}_t, \boldsymbol{p(t)}) \tag{5a}$$

$$\text{subject to:} \quad \mathbb{E}\left[\boldsymbol{B}_{t+1} = \boldsymbol{B}_t + \boldsymbol{U}_t + \boldsymbol{W}_t\right] \tag{5b}$$

$$d_{kn} \leq b_{kn} \leq B_{max}, \tag{5c}$$

$$0 \leq u_{kn} \leq B_{max}, \tag{5d}$$

$$\text{with: } k = t, t+1, \ldots, t+M-1\,.$$

Constraint (5c) defines the energy buffer limitations, ensuring through d_{kn} that the energy buffer of BS n contains enough energy to fulfill the expected demand. Constraint (5d) defines the limits in the energy purchasing process. Since the optimization problem must be solved at runtime, it is strongly preferable to choose a convex optimization formulation such as Eq. (5). Here, we have used the CVX tool [19] to obtain the optimal solution $\boldsymbol{U}_t^* = [u_{kn}^*]$, stating the optimal amount of energy that each BS n shall purchase in time slot $k = t, \ldots, t+M-1$.

Optimization algorithm: the MPC controller performs as follows [20]:

1) **Step 1:** at the beginning of time slot t, the system state is obtained, that is: energy buffer levels for all BSs, the harvested energy, traffic load forecasts and energy price for the next M hours (the optimization horizon).
2) **Step 2:** the control problem of Eq. (5) is solved, yielding a sequence of control actions over the time horizon (M slots).

TABLE 1: System parameters used in the numerical results.

Parameter	Value
Number of BSs, N	25
Number of BS traffic load clusters	5
Minimum BS energy consumption, P_0	13.6 W
Parameter α	1.1
Energy buffer capacity, $B_{\max}$	100 W h
Reference energy threshold, B_{ref}	100 W h
MPC optimization horizon M	24 h
Time slot duration, Δ	1 h

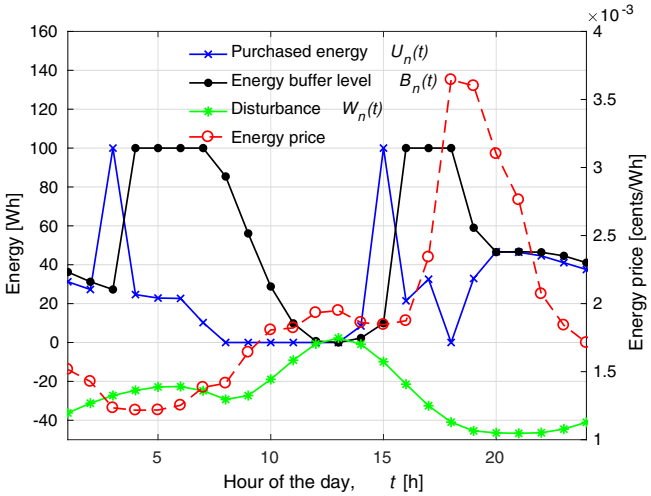


Fig. 4: MPC behavior over one day.

3) **Step 3:** only the first control action is performed and the system state is updated upon implementing the required energy purchases.
4) **Step 4:** at the next time step $t+1$, forecasts are updated and the optimization cycle is repeated from **Step 1**.

IV. Performance Evaluation

In this section, we discuss some selected numerical results for the scenario of Sections II and III. The parameters that were used for the simulations are listed in Table 1. Our proposed energy management framework, termed **MPC**, is compared against a benchmark strategy, termed **B1**. With **B1**, each BS n at time slot t purchases the amount of energy $\max(0, B_{\text{ref}} - B_n(t))$, so that its EB levels remains as much as possible equal to B_{ref}. This approach is not taking into account future system states such as traffic load, harvested energy and energy price, unlike **MPC** does.

In Fig. 4, we show the system operation of the **MPC** approach for a certain BS n. Based on the energy price $p(t)$ and on the expected behavior of $\mathbb{E}[H_n(t) - O_n(t)]$, the adaptive control block purchases energy over time. Two peaks in the purchase process (blue curve) can be observed: one is around four in the morning ($t = 4$), due to the low energy price at that time. The other one is around $t = 16$ (4 pm) when $\mathbb{E}[H_n(t) - O_n(t)]$ (green curve) decreases, which corresponds to an increase in the expected traffic load and to a corresponding decrease in the expected harvested energy.

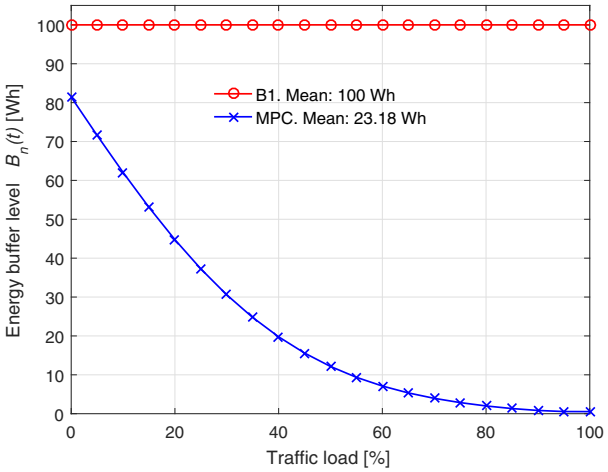


Fig. 5: Average EB level *vs* traffic load [%] in the system.

A comparison of the average BS EB level is presented in Fig. 5 as a function of the traffic load. In **B1**, all BSs maintain the EB level completely full for every traffic load, since B_{ref} has been set equal to $B_{\max}$. On the other hand, EB levels for **MPC** decrease with an increasing load, due to the fact that only the energy that is strictly required to serve the expected demand is purchased at each time slot t.

It should be noted that also with **MPC** the EB levels can be kept above a certain minimum threshold $B_{\min} > 0$, if needed. To achieve this, it is sufficient to replace d_{kn} with $\max(d_{kn}, B_{\min})$ in constraint (5c). In Fig. 5, we consider that there is no need to have extra stored energy, as long as the traffic demand is fully served. With MPC, no BS can run out of service and, if threshold $B_{\min}$ is set to any value greater than zero, then energy will be purchased as soon as any BS buffer level decreases below it.

The amount of purchased energy is compared between both schemes across different traffic load configurations in Fig. 6. As expected, the higher the traffic load, the higher the amount of energy that is purchased from the power grid. However, **MPC** leads to a reduction in the purchased energy of about 24% on average. Finally, a comparison of the energy cost (measured in $ over watts-hours) of the two schemes is shown in Fig. 7. As in the previous plot, the higher the traffic load, the lower the gap between the two schemes, as more energy is to be purchased. The biggest savings are achieved when the traffic load is in the range $[20-60]\%$ and **MPC** provides cost savings close to 20%, on average. In conclusion, the use of the proposed approach is beneficial for two reasons: (i) it maintains the EB levels as low as possible, while still serving all users and minimizing the energy losses due to an inefficient utilization of the energy storage; and (ii) it reduces the amount of energy that the BSs purchase from the electrical grid.

V. Conclusions

The management of sustainable mobile networks has been tackled in this work. We advocate base stations equipped with energy harvesting and storage capabilities, where each

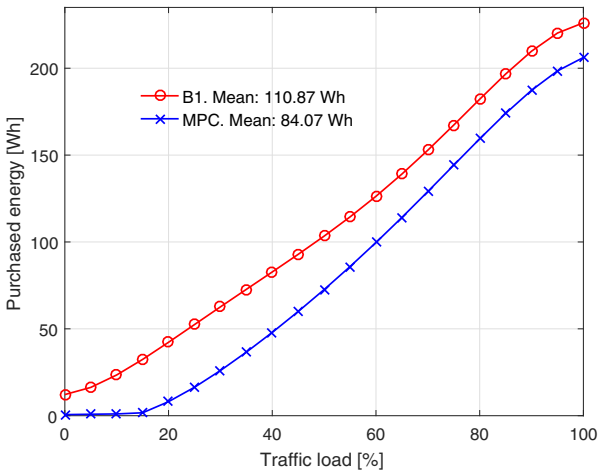


Fig. 6: Purchased energy *vs* traffic load [%] in the system.

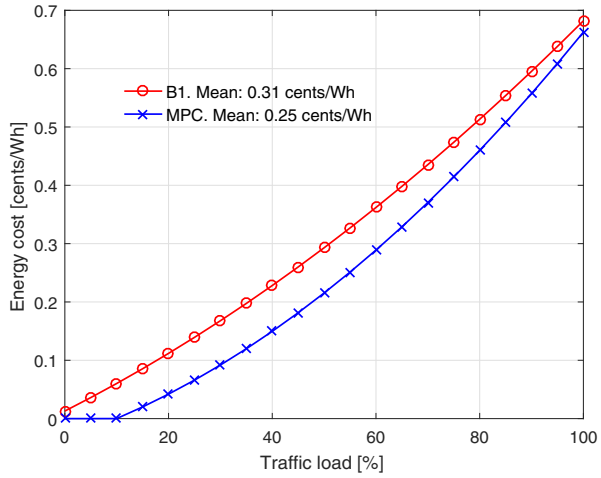


Fig. 7: Energy cost *vs* traffic load [%] in the system.

of them can acquire energy from the environment, use it to serve its local traffic or keep it in a local energy storage for later use. Whenever the harvested energy is insufficient to serve the traffic demand, BSs can purchase energy from the power grid. Within this setting, a management strategy has been proposed to reduce the amount of energy that is purchased from the power grid. This strategy controls the amount of purchased energy from the electrical grid over time, considering the amount of energy harvested, the traffic load, and hourly energy prices (set by the power grid operator). This framework combines pattern forecasting and adaptive control. In the forecasting block, the harvested energy and traffic load processes are predicted through an LSTM neural network, allowing each BS to independently predict its behavior over time. Hence, these forecasts are fed into an adaptive control block, where foresighted optimization is performed through MPC. Numerical results, obtained with real-world traces, show cost savings close to 20% and reductions in the amount of purchased energy from the electrical grid of about 24%, when compared to the case where future system states are not taken into account.

ACKNOWLEDGMENT

This work has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 675891 (SCAVENGE).

REFERENCES

[1] ICT Footprint - European Framework Initiative for Energy & Environmental Efficiency in the ICT Sector. [Online]. Available: https://ictfootprint.eu/es/about/ict-carbon-footprint/ict-carbon-footprint

[2] A. S. Andrae and T. Edler, "On global electricity usage of communication technology: trends to 2030," *Challenges*, vol. 6, no. 1, pp. 117–157, April 2015.

[3] GSMA - The Mobile Economy 2018 report. [Online]. Available: https://www.gsma.com/mobileeconomy/

[4] D. Niyato, X. Lu, and P. Wang, "Adaptive power management for wireless base stations in a smart grid environment," *IEEE Wireless Communications*, vol. 19, no. 6, 2012.

[5] X. Huang, T. Han, and N. Ansari, "Smart grid enabled mobile networks: Jointly optimizing BS operation and power distribution," *IEEE/ACM Transactions on Networking*, 2017.

[6] T. Han and N. Ansari, "On optimizing green energy utilization for cellular networks with hybrid energy supplies," *IEEE Transactions on Wireless Communications*, vol. 12, no. 8, pp. 3872–3882, 2013.

[7] J. Leithon, T. J. Lim, and S. Sun, "Energy exchange among base stations in a cellular network through the smart grid," in *IEEE International Conference on Communications (ICC)*. Sydney, Australia: IEEE, June 2014, pp. 4036–4041.

[8] M. Miozzo, D. Zordan, P. Dini, and M. Rossi, "SolarStat: Modeling photovoltaic sources through stochastic markov processes," in *IEEE International Energy Conference (ENERGYCON)*. Cavtat, Croatia: IEEE, May 2014, pp. 688–695.

[9] Telecom Italia (TIM) - Open Big Data Challenge 2015. [Online]. Available: https://dandelion.eu/datamine/open-big-data/

[10] D. Pelleg, A. W. Moore *et al.*, "X-means: Extending K-means with efficient estimation of the number of clusters," in *Proceedings of the Seventeenth International Conference on Machine Learning (ICML)*, vol. 1, San Francisco, CA, USA, June 2000, pp. 727–734.

[11] D. Zordan, M. Miozzo, P. Dini, and M. Rossi, "When telecommunications networks meet energy grids: cellular networks with energy harvesting and trading capabilities," *IEEE Communications Magazine*, vol. 53, no. 6, pp. 117–123, June 2015.

[12] US National Grid - Electric Service. [Online]. Available: https://www9.nationalgridus.com/niagaramohawk/business/rates/3_elec.asp

[13] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep learning*. MIT press Cambridge, 2016, vol. 1.

[14] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov 1997.

[15] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *International Conference on Learning Representations (ICLR)*. San Diego, CA, USA: IEEE, May 2015.

[16] F. Chollet *et al.*, "Keras," https://github.com/keras-team/keras, 2015.

[17] M. Abadi *et al.*, "TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems," 2015, software available from tensorflow.org. [Online]. Available: https://www.tensorflow.org/

[18] J. M. Maciejowski and X. Yang, "Fault tolerant control using Gaussian processes and model predictive control," in *Control and Fault-Tolerant Systems (SysTol), 2013 Conference on*. IEEE, 2013, pp. 1–12.

[19] M. Grant, S. Boyd, and Y. Ye. (2008) CVX: Matlab software for disciplined convex programming. [Online]. Available: http://cvxr.com/cvx/

[20] T. Zhang, Y. Zhang, H. Lei, B. Guo, and Y. Zha, "Optimal microgrid operation based on model predictive control framework," in *3rd International Conference on Control, Automation and Robotics (ICCAR)*. Nagoya, Japan: IEEE, Apr 2017, pp. 575–581.